

国家科学图书馆青年人才领域前沿项目

基于开放会议资源描述的 **RDA** 和 **DC** 比较研究

项目负责人：柴苗岭

课题组成员：徐源，韩红，刘春江

2013年9月

文摘：本文以中国科学院国家科学图书馆成都分馆牵头完成的“重要会议资源采集与服务系统”的会议资源为例，对 RDA 与 DC 描述资源进入了深入的比较研究。首先，本文建立了基于 RDA 的会议资源元数据描述方案，完成了会议资源 DC 元数据到 RDA 元数据的映射。通过实践应用的比较发现，DC 和 RDA 在语义关系上存在着三种对应关系。在资源描述上，DC 以其专有性在网络资源的描述上仍占有一定优势；而 RDA 以其深厚的图书馆基础，在资源描述领域上有着较强优势，但在其它社区的交换上仍存在着一系列问题，目前还无法代替 DC 在其它社区中扮演的角色。

关键词：开放会议资源 DC RDA 语义 映射

资源描述与检索 (Resource Description and Access, RDA) 是近年来为适应新技术和新资源发布需要而出现的新的内容描述规范。作为新的超越各类元数据的资源描述和检索方式, RDA 的出现引起了信息组织领域的高度关注。

在中国目前 RDA 的研究集中与 MARC 之间的互映射关系上的研究, 截止 2013 年 1 月 7 日, 还没有对 DC 到 RDA 的映射进行研究的文献公开发表。本文力图从理论和实践两个方面, 尤其是通过对 RDA 与都柏林元数据 (Dublin Core, DC) 在重要会议开放资源描述中的实证研究, 在比较分析中发现 RDA 的优势和不足。

为未来中国科学院国家科学图书馆、在数字资源的描述方法上采用何种元数据描述、对采用 DC 描述的资源与 RDA 进行映射等做出了探讨。

同时, 本文的 RDA 会议资源描述方案将成为重要会议开放资源采集与服务平台的新的元数据方案, 会议资源描述方案的 DC 到 RDA 的映射表将成为该平台与其它社区交换资源时的重要依据。

本文初步总结了二者在资源描述效用上的差异, 为进一步推广和优化 RDA 应用具有重要的启发作用。

基于重要会议开放资源描述的 RDA 和 DC 比较研究

中国科学院国家科学图书馆成都分馆于 2009 年启动了“重要会议开放资源采集与服务系统”项目作为其数图先期启动项目。该项目主要是针对网络上的重要会议开放资源进行的采集和服务。本部分的 RDA 和 DC 描述实践主要针对该采集服务平台上的资源进行, 制定了 RDA 资源描述集, 在原有成果上将三个 DC 的资源描述集修改为了两个描述集, 并且进行了 DC 到 RDA 的映射描述。

1. 引言

1.1 发展现状

如何帮助用户发现需求, 满足用户需求一直以来是图书馆学研究的重要内容。随着信息环境的变化, 人类的知识呈几何数增长, 用户的需求转向对知识的需求, 从而使图书馆传统的资源揭示内容和方式受到了严峻的挑战, 决定了图书馆的信息组织向深层次揭示信息内容和建立信息资源关联等方向发展。因此, 元数据作为图书馆信息资源描述的方式决定了资源描述的深度和关联程度。

从图书馆常用的资源描述元数据方案中来看, 自 1978 年就开始使用的, 经多次修定后的 AACR2 (Anglo-American Cataloguing Rules: Second Edition, 英美编目条例第 2 版) 是目前英美地区资源描述的主要标准, 也是我国《西文文献著录条例》的基础。它与 ISBD (International Standard Bibliographic Description, 国际标准书目著录) 相配套, 规定了文献著录的数据内容和数据呈现形式。就内容的揭示情况来说, AACR2 范围广, 内容详细, 注重载体的描述, 但不擅长数字文献和不同载体的文献描述。近年来, AACR2 针对电子资源和关联数据进行了适当的修

改，但在其指导下的书目数据，如 MARC (Machine-Readable Cataloging, 机读目录) 数据仍缺少完善的逻辑结构。这就导致了资源在分解成元素转换成关联数据时，因为关联度不够而存在一定的缺陷。

1995 年，旨在对网上各种电子信息资源进行描述的 DC 的出现满足了互联网下的资源语义信息描述的需要。该元数据由 15 个核心元素以及扩展元素、元素修饰词、编码体系修饰词、抽象模型、应用纲要、规范等组成。但随着 DC 的发展，目前更加吸收了本体、关联数据的理论、方法，有些老的方法在废弃，如修饰词不用了，Warwick 框架也没有真正应用，但 DC 采用的属性、子元素、抽象模型、新加坡框架，很多都是受语义网技术的影响下改变。因为 DC 具有简练、易于理解、可扩展性强的特点，在与其它元数据进行桥接时更显示其方便。DC 的出现较好地解决网络资源的发现、控制和管理问题，但它的语义关系原理是根据层次关系组织概念，不能用于指定关系的特征语义关系，在语义揭示上仍然有一定的局限。

随着 FRBR (Functional Requirements for Bibliographic Records, 书目记录规范功能需求) 概念的提出，AACR2 决定对其内容升级，修改过时的内容，开始制定 AACR3 (Anglo-American Cataloguing Rules: Third Edition, 英美编目条例第 3 版)。在修订过程中，JSC (Joint Steering Committee for the Revision of AACR, 修订工作联合指导委员会) 删除了冗余术语，修订了无法跟上数字时代的资源类型的分类方式和一般资料标识 (General Material Designation, 简称 GMD) 等。但后来发现修改内容过多，于是在 2005 年 4 月，在芝加哥举行的 JSC 和 CoP (The Committee of Principals, AACR 原则委员会) 会议上，正式把 AACR3 命名为 Resource Description & Access (资源描述与检索，简称 RDA)¹。

在历经多次修改后，RDA 于 2010 年 6 月正式发布。它遵照 2009 新的 ICP (Statement of International Cataloging Principles) 原则设计，把自己定位为一个元数据内容标准，不像 AACR2 那样还要指导具体的格式规范。

RDA 基于“实体-关系模型”，继承了 FRBR 和规范数据的功能需求 (Functional Requirements for Authority Data, 简称 FRAD)² 中对实体、属性、关系和用户任务的定义，更注重作品和内容的表达，在作品的描述、关系和检索点的规范。RDA 的这些特性使得它在对数字化资源的描述和各种资源类型进行控制时有较好的表现。在某种意义上，RDA 可以说是图书馆数据的首次网络语义表示标准。

¹ 张秀兰. 书目描述与检索的最新内容标准. [J] 图书情报工作. 2006 (3): 95-97

² IFLA. 书目记录的功能需求中文版 [EB/OL]. [2012-02-11]. <http://archive.ifla.org/VII/s13/frbr/frbr-zh.pdf>

1.2 研究现状

1.2.1 国内研究现状

(1)RDA 的相关研究

A.RDA 的相关文献研究

以 RDA、“资源描述与检索”为关键词在 CNKI（Chinese National Knowledge Infrastructure，国家知识基础设施）和 VIP（VIP Database for Chinese Technical Periodicals，维普中文科技期刊数据库）中进行检索。剔除无关、重复数据,共查到 RDA 相关文献 38 篇，检索时间为 2012 年 2 月 20 日。检索结果见表 1.1。

表1.1 RDA国内研究文献数量

时间	2006	2007	2008	2009	2010	2011
篇数	2	7	3	1	10	15

从研究的内容上来看，可以分为若干情况，如表 1.2。

表1.2 RDA国内研究主题分布

内容	篇数
与 AACR2、编目规则等的比较	3
RDA 测试结果、内容翻译	9
修订及发展历程	5
内容特征	2
概述、评价及影响	9
RDA 产生的影响及前景	7
应用	3

由上面 2 个表的统计结果来看，文献的发表时段从 2006 年 3 月到 2011 年 12 月，距离 RDA 的首次确定有一年左右的时滞。RDA 的研究在 2007 年有一个小高潮，这与 RDA 的内容推出有关系。但在其后的 2 年，国内对 RDA 的关注度有所降低。2009 年，只有 1 篇相关文献。在 2010 年 RDA 进入国家级图书馆的测试阶段后，国内对 RDA 的研究再次进入高潮，特别是 2011 年，图书馆界受到三大国家级图书馆的 RDA 测试及未来将使用 RDA 的声明鼓励,加强了对 RDA 的关注。

从相关文献的内容描述情况来看，与 RDA 相关的文献主要集中在其内容的翻译、介绍和评价上，各有 9 篇文献；其次是 RDA 产生的影响及前景，约有 7

篇；第三是修订及发展历程，有 5 篇；排名第四的分别是各有 3 篇文献的与编目相关规则、规范比较和 RDA 应用，最后是内容特征的文献，有 2 篇。从文献发表数量的内容分布情况，不难看出，国内目前对 RDA 的研究者主要还是集中在图书馆领域，对 RDA 研究的也以 RDA 发布内容介绍上为主，对 RDA 的实际应用、逻辑框架的解析较少。

值得一提的是关于 RDA 应用研究的 2 篇文献，它们分别是 2010 年 4、5 月由宋登汉发表的《基于 RDA 的中国古籍版本资源描述设计》^{3,4}和《图书馆 OPAC 的 FRBR 实践及相关思考——来自 RDA 标准的启示》⁵。其中，武汉大学的宋登汉等针对中国古典文献流传中产生的三种版本特性，就中国古籍版本资源就进行了分析，设计出了 RDA 元素集，是 RDA 本地化的一次尝试，也是目前公开发表的唯一一篇基于 RDA 元素应用的文章。

B.国内研究人员及项目

近年来，国内对 RDA 的研究较多的有华东师范大学图书馆的胡小菁，国家图书馆顾犇、高红，上海交大图书馆的王绍平等。他们分别对 RDA 与网络技术的关联、与编目规范的关系，及与 MARC 字段的映射进行了研究。如上海图书馆的刘炜、华东师范的胡小菁等对 AACR2 下的编目进行分析，并讨论了 RDA 化编目的书目数据用关联数据发布的可能性⁶。

在 RDA 与 MARC 的映射上，中国科学院国家科学图书馆于 2011 年对 RDA 到 CNMARC 的映射进行了尝试，对 CNMARC 从规范数据和书目数据两方面做了适应 RDA 的字段增删修改，然后结合了中国科学院国家科学图书馆的“联合目录集成服务系统”进行实测，发现虽然修改后的格式不会有元数据不兼容的情况，但会给用户在使用上增加语义上的理解难度。

在 RDA 的本地化研究方面，华东师范大学图书馆胡小菁申请的社会科学基金项目《<资源描述与检索>的中文化及其应用研究》⁷，于 2011 年 6 月通过了立项。该项目的通过开始了 RDA 中文化的探索，重点对 RDA 的 CNMARC 化进行相关研究。

(2) DC 与 RDA 关系的研究现状

在 DC 与 RDA 之间的互操作研究上，作者从 CNKI 和 VIP 的全文期刊数据库中使用“DC 元数据/都柏林元数据”+“RDA/资源描述与检索”组合检索，没有查询到相关的研究论文和学位论文。

³宋登汉, 周迪, 李明杰. 基于 RDA 的中国古籍版本资源描述设计 (一) [J]. 图书馆. 2010(4):51-53

⁴宋登汉, 周迪, 李明杰. 基于 RDA 的中国古籍版本资源描述设计 (二) [J]. 图书馆. 2010(5):49-52

⁵高红, 靖翠峥. 图书馆 OPAC 的 FRBR 实践及相关思考——来自 RDA 标准的启示[J].国家图书馆学刊 2011(2):23-29

⁶刘炜, 胡小菁, 钱国富, 张春景, 夏翠娟 RDA 与关联数据 [J]中国图书馆学报.2012(1):36-44

⁷ 编目精灵 (胡小菁),《资源描述与检索》的中文化及其应用研究 [EB/OL].[2012-01-05]http://catwizard.net/posts/20110617193300.html

以“DC 元数据/都柏林元数据”为关键词进行题名检索，在 CNKI 中查询到相关论文 75 篇，研究时间范围在 2002 年-2012 年，研究内容涉及以下四点：

第一，DC 自身内容研究，如 DC 模型和 DC 发展综述。从发表的文献和相关研究成果的发表情况来看，2005 年以前，我国对 DC 和 DCMI 有较多的关注和合作，如上海图书馆以吴建中、刘炜⁸为代表的对 DC 的全面介绍和引进，并翻译了 DC 的中文版内容。但 2005 年以后随 DC 理论的成型，对 DC 理论介绍的逐渐变少。以刘娟在《2005 年以来我国 DC 元数据研究综述》⁹中对 DC 在 2005 年后的变化进行了综述。

第二，DC 与其它元数据的互操作研究，如 DC 与 MODS，DC 与 MARC 的互操作对比研究，伯琼等在 DC、MARC21 和 CNMARC 的映射及匹配上的研究[7]等。

第三，DC 的应用研究方面，除在图书馆领域里的使用和探索外，DC 在非图书馆机构中也展露头脚，如《农业网站元数据管理信息系统的研究与实现》¹⁰中，作者潘冲把 DC 元数据应用到了农业数据库的管理，包得海，刘昉等把 DC 也放入到旅游数据¹¹的管理中应用。

第四，DC 与语义网的关联。例如，如何语义 Web 上 DC 元数据的描述方法的研究、自动从语义网上抓取数据的方法途径等。

从目前国内文献的调研情况来看，RDA 还属于新生事物，国内的图书馆界对其认知度不够，从网络调查“国内 RDA 认知度调查”¹²的结果来看，对 RDA 规则较熟悉和熟悉的人数只占参与调查人数的 14.44%，信息组织人员对 RDA 总体来说认知度不高，也就无法深入到语义级的对比。这也是在文献检索中没有发现与 RDA 与 DC 的互映射研究相关论文的原因。

1.2.2 国外研究现状

(1) RDA 研究现状

从目前国际上对 RDA 的研究来看，目前主要集中在三个方向的研究：

第一，RDA 与不同元数据间的互描述，如 MARC21 字段中的载体类型、内容类型、介质类型所对应的 RDA 描述，DC 元数据与 RDA 之间的描述。

第二，RDA 与 ISBD、ISSN 等标准之间的兼容。如 RDA ISBD0 与 RDA 载体类型、内容类型的映射，ISBD 与 RDA 元素的映射。如 ISBD0 与 RDA 载体类型、内容类型的映射，ISBD 与 RDA 元素的映射。映射完成后最终以 RDF/SKOS 格式置于开放元数据注册（Open Metadata Registry）。

⁸吴建中，DC 元数据.[M]. 上海：上海科学技术文献出版社，2004.

⁹刘娟[J]《新世纪图书馆》2011 年（8）：37-41

¹⁰潘冲. 农业网站元数据管理信息系统的研究与实现.[D]中国农业科学院

¹¹包得海，刘昉. DC 元数据在旅游资源数据库中的应用研究[J].现代计算机(专业版) 2011(4):15-17

¹² RDA 认知调查正式版 [EB/OL].[2012-05-02] <http://www.sojump.com/viewstat/1286336.aspx>

第三，RDA 与开源软件的合作。如 eXtensible Catalog (XC)作为开源发现软件放入 RDAtoolkit 平台中，作为 RDA 数据转化的官方软件。

随着近年来 RDA 内容的日趋完善，RDA 的研究也逐步从理论走向应用。许多国家级图书馆纷纷进行了测试。美国国会图书馆提出了“书目框架转变行动”¹³，已于 2011 年 11 月正式开始 RDA 编目。由于 RDA 作为新生事物，在内容上还不够完善，各国图书馆在对 RDA 接受的同时，也保持着谨慎、观望的态度。

RDA 这种“高级标准”是否如其所说的能够跨越载体和各类元数据完美表达资源中的语义关系，在描述中相比其它元数据有什么样的优势，存在什么样的问题，尚需要通过实际应用得出结论。

(2) RDA 与 DC 互映射研究现状

RDA 与 DC 的官方映射研究由美国雪城大学的 Diane Hillmann(黛安·何尔曼)和斯特拉斯克莱德大学 Gordon Dunsire(戈登·邓塞尔)领导的 DCMI/RDA Task Group (DCMI/RDA 工作组)承担。该工作组主要从三个方面进行映射研究¹⁴：

第一，RDA Element Vocabulary (RDA 元素词)的研究¹⁵。目前，在 2007 年版的 427 个 RDA 元素表的基础上，完成了 RDA 的三层元数据：Elements (核心元素)、Element sub-types (元素子类型)和 Sub-Element (子元素)的分析，以及 RDA TOOLKIT 集成包的翻译。

第二，开发基于 RDA 的 FRBR 和 FRAD 模型上的 DC 应用纲要¹⁶。完成了用 RDF 对 RDA 模型实体进行定义，包括了标签、定义和注释范围。已完成了 RDA 第一组实体的研究

第三，确定 RDA Value Vocabularies (RDA 值)在 RDFS 或 SKOS 中的定义上¹⁷，通过 ISBD 分析了 DC 元素和 RDA 的类和属性上的关系。

该工作组的 Diane Hillmann 制定了 Draft mapping RDA to Dublin Core (RDA 到 DC 元数据的映射草本)¹⁸。但是，他发现了 RDA 中 2.3.7 章反应属性、个体和表达的元素在 DC 中没有对应的表达，已被映射到 DC 的 TYPE (类型)元素中的音乐内容表现形式在 RDA 中没有对应的词汇编码方案，其表述是否有效等问题。

¹³ Transforming our Bibliographic Framework: A Statement from the Library of Congress (May 13, 2011)[EB/OL].[2012-01-05]

¹⁴ Workplan for RDA Vocabularies Project.[EB/OL].[2012-01-05]<http://dublincore.org/dcmirdataskgroup/>

¹⁵ Task 1.1 Analysis of RDA data elements.

[EB/OL].[2012-04-06]<http://dublincore.org/dcmirdataskgroup/analysisTask1>

¹⁶ Develop a DC Application Profile for RDA based on FRBR and FRAD. [EB/OL].[2012-04-06]

<http://dublincore.org/dcmirdataskgroup/>

¹⁷ Identify in-line value vocabularies as candidates for publication in RDFS or SKOS. [EB/OL].[2012-04-06]

<http://dublincore.org/dcmirdataskgroup/>

¹⁸ Danskin, Alan. (2010, February 22). G'day and question regarding RDA mapping to DC. Retrieved from the DC-RDA Archives:[EB/OL].[2012-05-05]

<https://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind1002&L=DC-RDA&F=&S=&P=1293>

此后，哥伦比亚大学，芝加哥大学和伊利诺大学香槟分校也根据该草本尝试了 Dublin Core Element（都柏林核心元数据）到 RDA 的映射测试，对 RDA 与 DC 的语义映射，角色、类型和元素语境，本地模式的发展以及与 FRBR 的实体关系进行了分析，总结了在映射过程存在的问题。¹⁹

DC 与 RDA 的研究主要在 DC 到 RDA 的单向映射上，但 RDA 还在发展中，不断有新的元素产生，映射关系的处理上还有待进一步研究。

1.3 研究设计

1.3.1 研究目标

RDA 期望成为一个全球范围内的大而全的标准格式，不仅能适用于英语国家和西方国家，而且可以作为世界范围内的资源描述与检索的内容标准；不仅能适用于传统文献类型的著录，也能适用于电子资源，从而覆盖所有资料类型属性的描述。但是，RDA 在数字图书馆环境下的资源描述中是否像它自己所说的那样，能全面体现资源间的关系，较之其它元数据更能表现语义关系呢？

本文希望通过对 RDA 和 DC 在资源描述效果上的比较研究，尤其是以重要会议开放资源采集与服务系统中的会议资源为例，建立 RDA 元素集，实现 DC 到 RDA 的映射，在描述实践上对比分析两种元数据，找出 RDA 在资源描述时存在的优势和不足。

1.3.2 研究思路

本文在简要介绍 RDA 和 DC 的基础上，首先从逻辑关系、理论来源及其自身结构等方面，对比 RDA、DC 两种元数据，然后以重要会议开放资源为例，通过建立 RDA 资源描述集、形成 MARC 与 RDA 的互映射表、DC 与 RDA 的互映射表等，分析资源描述完整性及元数据的互操作性，总结 RDA 在资源语义描述上、元数据互操作上、资源的综合集成上的优势和局限性。

1.3.3 研究方法

本文主要运用以下两种研究方法：

第一，比较分析法：从理论来源、理论实践两方面，对 RDA 与 DC 的描述能力进行对比。

第二，实证分析法：基于重要会议开放资源采集与服务系 RDA 和 DC 的设计实践。

¹⁹ Testing Resource Description and Access (RDA) with Dublin Core, Myung-Ja Han, Melanie Wacker, Judith Darrt.[EB/OL].[2012-02-15]

2. DC 和 RDA 简介

2.1 DC 元数据

DC 元数据即 Dublin Core Metadata (都柏林核心元数据),是一套用于网络资源描述和资源发现的描述性元数据。它产生于 20 世纪 90 年代中期。随互联网的蓬勃发展,互联网上产生了大量的虚拟文档,当时 OCLC (OCLC Online Computer Library Center, 联机图书馆中心)和 NCSA (National Center for Supercomputer Applications, 美国超级计算机应用中心)看到这类文档的发展前景。他们认为有必要对这部分文档进行管理,于是,1995 年两个机构联合在加拿大 Dublin (都柏林)召开了第一届元数据研讨会。在会议上,工作组确定了两个目标,其中之一就是“在描述网络资源核心元数据元素集上达成共识”²⁰。会后决定建立一个能对网上各种电子资源信息进行描述的、简单易操作的核心元数据元素集,即 Dublin 核心元数据。

DC 元数据从最早的 13 个核心元素发展到 15 个核心元素以及 DCTERM、词表,共计 55 个元素,成为一套由资源模型、描述集模型和词汇模型组成的 DC 抽象模型。这一发展过程经历了 17 年之久。

其后会议陆续召开,但都不固定,直到 2001 年,保持每年召开一次的频率,2008 年开始,DCMI 独立主持 DC 元数据学术研讨会。

从 DC 的发展历程上来看,可以分为三个阶段。

(1) 1995 年至 2000 年的基础理论完善期

1995 年、1996 年 4 月和 1996 年 9 月的三次元数据会议,确定了 DC 的核心元素,把从最早设计的 13 个增加到现在的 15 个元素的固定规模,并确定了其数据模型为 Warwick 框架的元数据结构等。

在 1997 至 2000 年的短短的三年里,DC 元数据会议总共召开了五次,先后确定了附加的 DC 修饰词,即堪培拉修饰词,推出了著名的关联元素和 1:1 原则等。这一系列的举措,搭建了 DC 的元数据框架,使 DC 在理论上更加完善起来。

(2) 2000 年-2007 年的实际应用期

2000 之后,DC 元数据研究会议基本上保持了一年开一次会议的频率。2000 年至 2003 年期间,共召开了三次会议。其逻辑模型结构也于 2005 年形成,并于 2007 年对该模型进行了升级,推出了 DC1.1。这一阶段是 DC 的成熟期。在 DC1.1 方案推出后,不断被研究和应用,标志着 DC 元数据真正走向了应用。

(3) 2008 年以后的全面应用,新方法和思路的吸收

自 2008 年,DCMI 不再依附于 OCLC,开始独立主持 DC 元数据学术研讨会。随着语义网和关联数据的兴起,DC 也把注意力转向了语义网、语义揭示和建立其它元数据的映射。DC 在其随后的发展过程中吸收了本体、关联数据的理论、方法等,受语义网技术的影响,DC

²⁰吴建中, DC 元数据.[M]. 上海: 上海科学技术文献出版社, 2004.

采用了属性、子元素、抽象模型、新加坡框架等，但它同时也摒弃了一些老的方法，如修饰词和 Warwick 框架都没有再使用。始终保持着旺盛的生命力。

2.2 RDA 简介

2.2.1 RDA 产生背景

在开放式、分布式的网络环境下，存在着大量异构的信息资源。这些异构资源不断产生并发生交换。从图书馆角度来说，图书馆数据与外界的交换一直受到元数据格式的局限，不能直接被使用。同时，不同的国家、不同机构采用不同的资源描述方式，影响了数据交换的效率。图书馆界原有的描述资源的规范 AACR2 主要是针对 20 世纪 70 年代流行的卡片目录格式上的编目规则，与现有的机读目录不相匹配，特别是在 FRBR 产生后，发现更加无法适应当前用户对资源的描述和检索的需求。于是，1998 年成立的 Joint Steering Committee for the Revision of AACR（修订英美编目规则联合指导委员会），决定改进原有编目规则，制定新的编目规则，这就是后来的 RDA。

2.2.2 RDA 组织机构

RDA 是由 The Joint Steering Committee for Development of RDA（资源描述与检索发展联合指导委员会，JSC）进行指导和编撰的。该机构原名为 Joint Steering Committee for the Revision of AACR，主要由美国图书馆协会、澳大利亚编目委员会、英国国家图书馆、加拿大编目委员会、皇家特许图书馆和情报专业学院（CILIP）和美国国会图书馆等 7 个英美主要编目机构的代表组成。

RDA 就是在这 7 个机构的共同努力下制定的。JSC 把任务分解到 7 个工作组^[11]，格式变化工作组主要负责对内容表达创建书目记录的可行性进行测试，GMD/SMD（General Material Designation/Standard Material Designation，资源类型标示/具体资料标识）工作组提出适合识别内容简介载体术语。这两个工作组已经分别于 2004 年和 2006 年完成了工作任务。目前还在进行中的小组有分别于 2005 年和 2006 年启动的两个样例小组，2005 年启动的 RDA 延伸小组，2006 年启动的附录工作小组，2007 年启动的 DDCMI 负责的 RDA 与 DC 映射小组，2008 年启动的 MARC 和 RDA 之间映射的工作组。MARC 与 RDA 在 RDA Toolkit 中已经推出了较成熟的互映射版本，而 DC 与 RDA 的映射草本虽然经过了美国三大高校的映射测试，但仍存在较多问题，没有推出正式版。

2.3 RDA 理论基础

1988 年，IFLA 顺应网络的出现，确定了基于“实体-关系模型（Entity-Relationship Model）的基于实体-关系模型（Entity-Relationship Model）

的 Functional Requirements for Bibliographic Records (书目功能记录需求, FRBR) 研究, FRBR 正是后来 RDA 的重要理论来源之一。

但是, 该研究只涉及了书目记录的功能需求研究, 没有讨论关于规范数据和主题规范的内容。应 FRBR 工作小组的要求, 后期又建立了 FRANAR 工作小组和 FRSAD 工作小组, 分别对“主题规范、叙词表、分类表的中心实体和这些实体之间的关系”²¹进行了后期研究, 2009 年和 2011 年正式发布了 Functional Requirements for Authority Data (规范数据功能, 简称 FRAD) 和 Functional Requirements for Subject Authority Data (主题规范数据的功能需, 简称 FRSAD)。“FR” 家族就此形成。

2.3.1 RDA 理论来源的分析

将三组规范的形成时间、分析技术、主要研究目标、实体、属性、关系及用户任务进行对比, 如下表。

表2.1 FRBR、FRAD、FRSAD内容比较

	FRBR	FRAD	FRSAD
最终形成时间	1998年终版, 2007年进行修改	2009年	2011年
分析技术	实体分析技术	实体分析技术	实体分析技术
主要研究目标	1.提供了一个结构化框架, 使书目记录中记录的数据与记录用户的需求之间产生关联 2.是推荐由国家书目机构创建的记录的基本功能级别	建立一个着力于数据本身的规范控制数据和规范数据国际共享功能需求的模型	1.就 FRBR 中第三组“关于”方面制定一个概念模型 2.就记录在主题规范记录中的数据和用户使用该记录的需求提供一个参考框架。 3实现图书馆界以及外界实现主题规范数据的国际分享与使用潜力的协助评估。

²¹ IFLA. 书目记录的功能需求中文版 [EB/OL].[2012-02-11]. <http://archive.ifla.org/VII/s13/frbr/frbr-zh.pdf>. P5

实体	<p>1.作品 (Work)、内容表达 (Expression)、载体表现 (Manifestation)、单件 (Item)</p> <p>2.个人 (Person)、团体 (Corporate Body)</p> <p>3.概念 (Concept)、物体 (Object)、事件 (Event)、地点 (Place)</p>	<p>1.作品 (Work)、内容表达(Expression)、载体表现 (Manifestation)、单件 (Item)</p> <p>2.个人 (Person)、家族 (Family)、团体 (Corporate Body)、名称 (Name)、标识符 (Identifier)、受控检索点 (Controlled Access Point)、规则 (Rules)、机构 (Agency)</p>	Thema、Nomen
属性描述对象	<p>实体内在属性和外部属性。描述了第一组实体的作品、内容表达、载体表现和单件,第二组实体的个人、团体,以及第三组实体的概念、实物、事件和地点的属性</p>	<p>包括个人、家族、团体、作品、内容表达、载体表现、单件、概念、实物、事件、地点、名称、标识符、受控检索点、规则和机构的属性</p>	Thema、Nomen
关系	<p>1.模型环境中的书目关系</p> <p>2.在高层次图表中描述的关系,包括三部分:作品、内容表达、载体表现和单件之间的关系,其次,个人与团体的关系,第三,主题关系</p> <p>3.第一组实体之间的其他关系。包括了作品到作品的关系,内容表达达到内容到达的关系,内容表达达到作品的关系,载体表现到载体表现的关系,载体表现到单件的关系,单件到单件的关系</p>	<p>1.模型环境中的规范关系,包含了四大类关系</p> <p>2.高层次图中描述的各种关系</p> <p>3.个人、家族、团体与作品间的关系,个人、家族、团体以及作品的变异名称之间的关系,受控检索点之间的关系</p>	<p>1.不同类型实体间的关系</p> <p>2.相同类型实体间的关系</p>
用户任务	<p>查找 (Find)、识别 (Identify)、选择 (Select)、获取 (Acquire)</p>	<p>查找 (Find)、识别 (Identify)、阐明关系 (contextualize)、提供依据 (Justify)</p>	<p>查找 (Find)、识别 (Identify)、选择 (Select)、发现 (Explore)</p>

将以上三组模型进行联系和对比,可以发现:

(1) 总体结构

FR 家族借鉴了软件工程领域的“E-R”实体关系模型。三组都采用了实体分析技术,即首先分析出实体,然后根据实体析出实体间的特征属性和关系。

（2）目标任务

FRBR 把书目记录中记录的数据与用户的需求之间建立联系，构建一套清晰定义的结构化框架。FRAD 则重于数据本身的规范控制数据和规范数据国际共享功能需求的模型建立，FRSAD 着力于提出主题规范记录中的数据和用户使用该记录需求的参考框架。三者互为补充。

（3）实体

FRBR 把实体划分出成了三组²²，第一组为书目记录命名或知识或艺术创造的产品，包括了作品、内容表达、载体表现及单件，这一组实体也是 FR 家族和 RDA 的实体。第二组为对知识或艺术内容物质产生与传播或其产品的保管负责的实体，包括个人和团体。第三组包括一系列附加的实体，它们作为知识或艺术创作的主题，主要由概念、实物、事件和地点四部分组成。

FRAD 的第一组实体与之相同，对第二组实体进行了修改和扩充，并增加了家族 (Family)、名称 (Name)、标识符 (Identifier)、受控检索点 (Controlled Access Point)、规则 (Rules)、机构 (Agency)。

FRSAD 则从 FRBR 的实体中抽象出 Thema、Nomen 两个实体，这两个名词来源于拉丁术语，本身没有预先存在的内容。Thema 可以看做“Subject (主题)”，“Topic (主题、标题)”和“Concept (概念)”，Nomen 是任何标志或标志序列，包括 FRAD 实体的 Name、Identifier 和 Controlled Access Point。

（4）属性

属性是关于实体的属性，包括了实体的内部属性和外部输入的属性。外部属性是指物理特征，如其一本图书的尺寸等载体情况，也包括了一些标识信息特性。外部属性则是分配给实体的一个主题说明或者背景。FRBR 和 FRAD 定义的属性都要求从用户角度来描述，所以都是逻辑层次的属性。简而言之，属性是通过一系列元素，把一个实体“作品”客观的描述出来，所列出来的元素越详尽，则越能帮助用户完成特定资源的查询。例如在会议录信息中，对“会议录”实体的物理属性和题名、形式、日期、等描述角度充分展现内容特征。

（5）实体关系

FRBR 表达了三个关系：

“E-R”实体模型环境中的书目关系，在高层次图表描述关系，即作品、内容表达、载体表现和单件之间的关系。

其次，第二组实体个人与团体的关系，主题关系的关系。

²² IFLA. 书目记录的功能需求中文版 [EB/OL].[2012-02-11]. <http://archive.ifla.org/VII/s13/frbr/frbr-zh.pdf>.
P17

第三，第一组实体之间的其他关系，包括了作品到作品的关系，内容表达到内容到达的关系，内容表达到作品的关系，载体表现到载体表现的关系，载体表现到单件的关系，单件到单件的关系。

FRAD 也包括了模型环境中的 4 大规范关系²³：

第一类关系代表一般层次的，实体-关系图中的实体类型间的关系，即作品、内容表达、载体表现和单件四者间的关系。

第二、三类关系代表包括规范记录中的，参见结构映射的关系（例如参照本身、信息附注、说明语或编目员附注）。

第四类是包括规范记录中嵌入的连接结构所反映的受控检索点实体的事例间的关系。

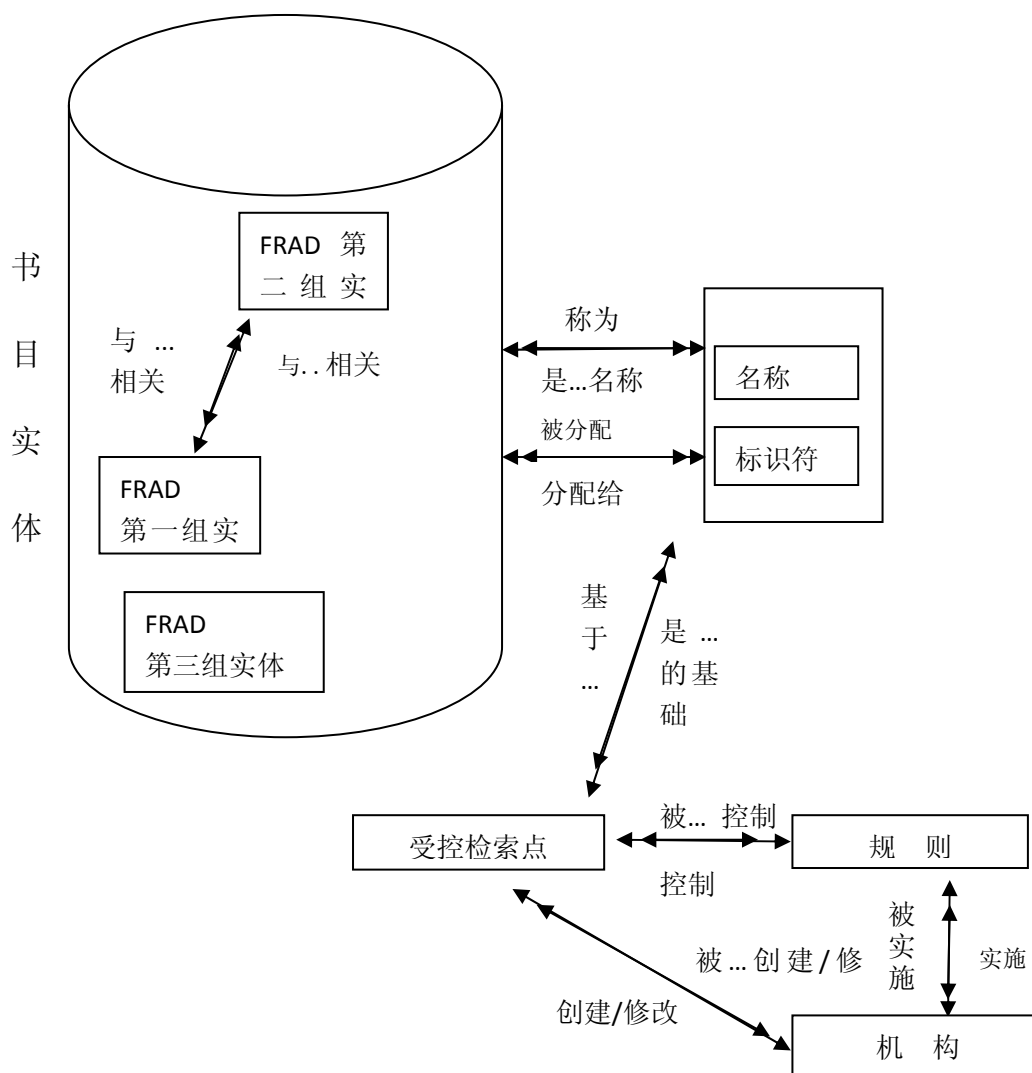


图2.2 规范数据的概念模型（本图根据FRAD模型修改而成²⁴）

²³ IFLA. 规范数据的功能需求概念模型 [EB/OL].[2012-02-15].
www.ifla.org/files/cataloguing/frad/frad_2009-zh.pdf. P50

FRSAD 模型则建立了两组关系。其中，一组是不同类型实体间的关系“作品到 THEMA”和“THEMA 到 NOMEN”的关系。另一组是相同类型实体间的关系，如“THEMA 到 THEMA”和“NOMEN 到 NOMEN”。需要提到的是，三组实体（三组实体主要是 FRBR 的实体，但在第二组实体中增加了 FRAD 的 Family）里面，每个实体都可以抽象出多个 Thema，每个 THEAM 都可以有多个 NOMEN，每个实体可有多个 THEMA，每个 HEMA 都可以对应一个 NOMEN，但如果 THEMA 为受控词或整合词时，则每个 THEMA 可以有多个 NOMEN，但每个 NOMEN 只能有一个 THEMA。关系图如下。

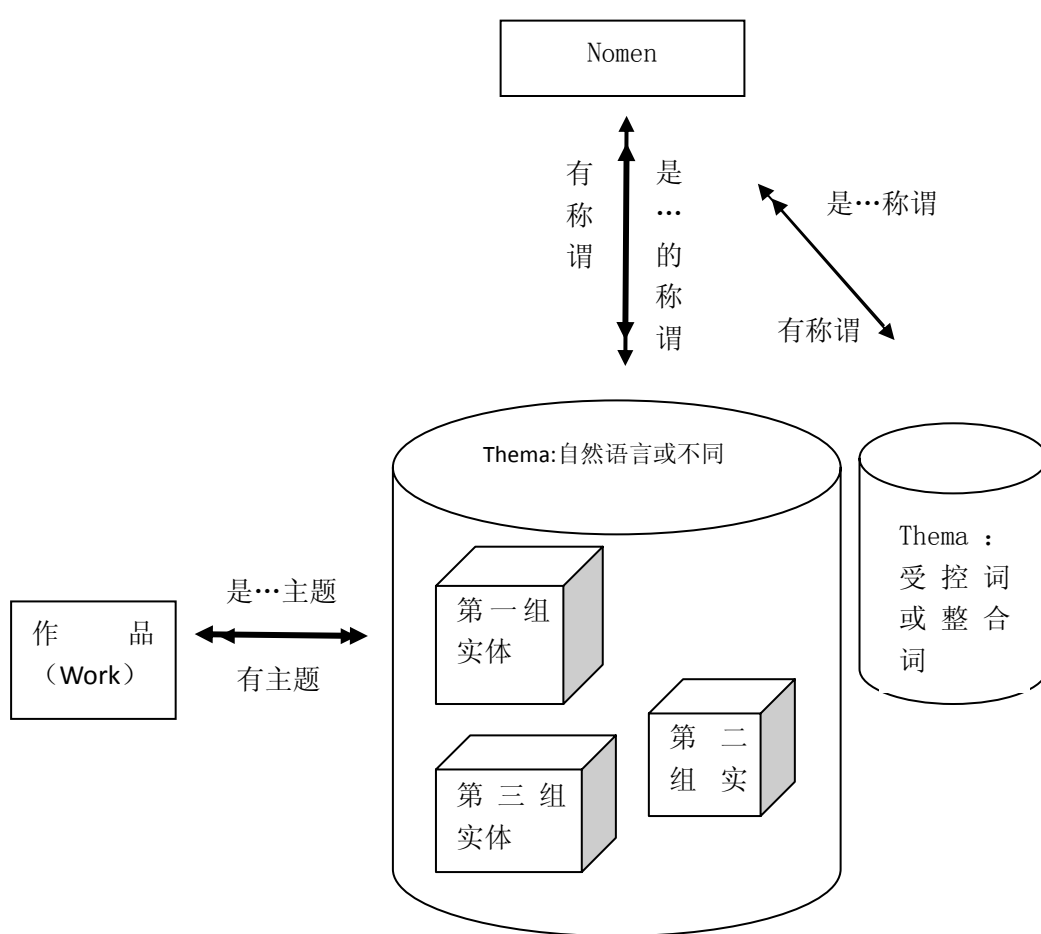


图2. 3: FRSAD的实体关系图, 本图为合成图²⁵

(6) 用户任务

²⁴ IFLA. 规范数据的功能需求概念模型 [EB/OL].[2012-02-15].

www.ifla.org/files/cataloguing/frad/frad_2009-zh.pdf.

²⁵ IFLA. Functional Requirements for Subject Authority Data (FRSAD): A Conceptual Model. [EB/OL].[2012-03-15]. <http://www.ifla.org/files/classification-and-indexing/functional-requirements-for-subject-authority-data/frsad-final-report.pdf>

用户可以分为系统用户和资源使用用户。

A.FRBR

FRBR 定义了四项基本任务，即查找（Find）实体、识别（Identify）实体、选择（Select）实体、获取（Acquire）实体。

这四个任务可以表述为：用户可以使用单一属性、属性的组合或实体间关系作为检索条件查找单一实体或一组实体；这一个实体能够被所描述的实体对应查找到，且两个或多个实体能被区识别；这一实体能在内容、物理形式等方面被用户选择取舍；一个实体可以通过购买、借阅或者以远程的方式获取或存取。

B.FRAD

FRAD 从书目数据角度出发，建立了查找（Find）实体、识别（Identify）实体、阐明关系（contextualize）和提供受控检索点基础的依据（Justify）四个任务。但 FRAD 没有效仿 FRBR 为用户任务进行评估分级。

与 FRBR 相比，FRAD 在 FRBR “查找”和“识别”两个任务的基础上，增加了使用属性及其关系在整个书目实体中能查找到相应实体和识别实体在受控检索点的名称形式两个任务。

FRAD 后两个任务则体现了其对关系的要求。如 FRAD 要求阐明 FRBR 定义的第二组实体，即个人、团体与作品三者间的关系。并提供规范记录创建者在选择名称或名称形式时建立检索点名称的依据。

C.FRSAD

用户使用主题规范数据，也定义了查找（Find）、识别（Identify）、选择（Select）、发现（Explore）一个或者多个主题实体四项任务。其中“发现（Explore）”任务是 FRSAD 提出的新任务，即能够发现主题和主题间称谓的关系。

另外三个任务查找、识别和选择任务分别来自于FRBR或者FRAD概念模型，只是任务内容变为“主题”。

2.3.2 “FR” 家族与 RDA 的关系

RDA 沿用了 FR 家庭的实体-分析技术和模型，并根据 FR 家族的内容进行了一些调整。

（1）实体

在实体部分，RDA 的元素集中吸收了 FRBR 的三组实体，和 FRAD 中的名称（Name），共同组成了 11 个实体。但在属性和关系的描述过程中，RDA 并未将“名称”单独列为实体阐述其属性和关系，而是将其放入 RDA 第 8 章个人、团体和家族的属性部分。

（2）实体间的属性

RDA 的实体属性可以分为对四部分实体的特征阐述，即作品和属性，内容表达和单件，个人、家族和团体的属性，概念、对象、事件和地点的属性。

(3) 实体间的关系

RDA 对 FRBR 中提到的三组关系，即：高层关系、主题关系和第一组实体外的其它关系，都不同程度的进行了继承。如 RDA 接受了 FRBR 高层关系中的第一组实体间的关系，但在“个人和团体”关系部分增加了“家族”，构成了个人、家族和团体关系。其次，将第二部分“作品”与其它实体的关系分解为：概念、对象、事件和地点的主要关系和其它关系。

从 FRAD 中引入的实体“名称”没有单独列为实体。

(4) 用户任务

RDA 用户任务的内容综合了 FRBR 和 FRAD 两部分，并扩大了两者用户任务的内涵。

RDA 从 FRBR 中吸收了发现 (Find)、识别 (Identify)、选择 (Select) 和获取 (Acquire) 四个用户任务。同时也吸收了 FRAD 中的发现 (Find)、识别 (Identify)、阐明关系 (Contextualize) 和提供依据 (Justify) 的内容，但将“阐明关系”更改为“澄清 (Clarify)”，把限定于个人、团体和作品之间的关系阐明扩大到所有实体之间的关系阐述。同时，把“提供依据”修改为“理解 (Understand)”，RDA 对 FRAD 中的提供名称为检索依据扩大为提供一套全面的指引和指令涵盖所有类型的内容和媒体。

3.1 重要会议开放资源概况

3.1.1 资源分类

第一，按收费情况区分，重要会议开放资源的基本类型可以分为免费 (Free)、开放获取 (OA) 和会员免费。

第二，从资源来源区分，重要会议开放资源可以分为会议录门户网站、会议网站、学会和机构网站、机构知识库 (IR) 以及 OA 出版物。

第三，从会议类型区分，重要会议开放资源包括国际性学术会议、国家级学术会议、机构级别的学术会议等。其中，以 Workshop 为最主要的会议类型。这类会议与 Conference (会议)、Symposium (座谈会)、Annual Meeting (年会) 相比，规模较小，形式较为非正式，其会议论文往往不正式发布，但比较前沿，往往是科研人员比较关注的，所以需要重点关注，也是本系统的建设重点。

第四，从会议的情况来看，会议的组织形式有一个或多个会议的历届会议集成在同一个网站上或者一个或多个会议的历届会议分散在不同网站上。

第五，网站提供的开放会议文献的文档形式中，绝大多数提供了会议文献全文，但有的网站仅提供了 PPT，还有的只提供了摘要；绝大多数提供单篇文献的链接，有些是以整本会议录的形式提供，甚至是将整本会议录压缩后提供；个别

会议提供音频（如 MP3）、视频文件。

第六，从资源描述的层次和深度来看，会议资源的描述包括了会议，会议又可分为单个会议和连续会议，其次是会议录和会议文献，共三个层次的描述。从组织的内容上来看，每个开放会议网站的会议内容有限，资源结构相对简单，揭示资源的层次和深度明显不够。

第七，从著录对象来看，重要会议开放资源可分为会议、会议录、会议文献。其中，会议是一个固定的机构组织的一个连续性的会议，也可能是由不同机构组成的连续性的会议。其次，存在着多级会议，即一个会议可能由个分会/卫星会议组成。第三，一个会议可以产生出多个会议录，一个会议录也有可能对应到多个会议上面。第四，一个会议录可能以开放获取或勉强的形式在网络上出现，也可能同时以常规出版和免费、开放共存的现象。

3.1.2 资源描述的主要内容

原有的重要会议开放资源描述内容由《重要开放会议元素名称、释义、描述规范》、《重要开放会议及会议录元素名称、释义、描述规范》、《重要开放会议文献元素名称、释义、描述规范》三部分组成，包括了会议信息、会议录信息和会议文献信息三部分，这都是资源描述的主要内容，如下：

（1）会议信息主要描述的内容

对多种途径发现和搜集到的会议资源站点进行简要描述，描述的内容包括以下内容。

表 4.1：会议信息主要描述的内容

基本描述字段		补充描述字段		
资源开放性	学科领域	资源发布者 (提供者)性质	资源 时 间范 围	资源 语种

1. OA 资源 2. 免费获取全文 3. 对会员、特定机构开放 4. 收费(订购) 5. 无会议全文 6. 多媒体开放资源 7. 非资源站点 8. 不确定	1. 自然科学, 并标注学科 2. 图书馆学情报学 3. 人文社会科学领域 4. 不确定	1. 重要机构 2. 一般机构 3. 不确定	英语, 德语, 法语
-----------------------------------------------------------------------------------------------------	-------------------------------------------------------	------------------------------	------------

上表中, 重要机构指国际性学(协)会、国家级学(协)会、重要学术与研发机构、重要资源门户等。

(2) 会议录信息描述内容

首先是基于开放获取会议资源站点一览表中的资源, 对开放会议资源站点的所有开放获取的会议进行描述, 描述字段如下表。

表 4. 2: 会议录信息主要描述的内容

字段		注:
基本描述字段	资源网址	
	会议录起始网址	
	会议名称	
	会议学科	整合中图法和科图法, 确定了会议学科分类表, 分类表共分 95 个类目
	会议地域范围	1. 国际性会议: 会议名称中标明为国际会议; 2. 两国会议: 会议名称中表表明为两国会议; 3. 发达国家国家性会议; 4. 发展中国家国家性会议; 5. 重要机构的机构会议; 6. 一般机构的机构会议。不在 1-5 范围内的会议。
	会议日期	yyyy 或 mm-dd-yyyy
开放内容	1. 会议全文 2. 会议陈述 3. 会议全文及会议陈述 4. 会议发言音频、视频	

	开放方式	1. OA 2. 免费 3. 对会员或内部用户开放
	开放程度	1. 所有论文开放 2. 部分论文开放
	二次文献数据库 收录情况	
	会议地点	国家,城市
	会议举办者	
	会议录题名	
	母体文献题名	
	会议录发布者	
	会议录发布日期	yyyy-mm-dd
	ISSN/ISBN	

(3) 会议文献描述内容

重要会议开放资源采集与服务系统中的会议文献主要来自于 PDF 和 TXT 版本的文字内容，目前主要描述内容如下。

表 4.3: 会议文献主要描述的内容

字段		注:	
基本描述 字段	资源网址		
	会议名称		
	作者信息	作者	
		作者机构	
		作者 EMAIL	
	摘要	1.Abstract 2.Summary 3.第一段段首	
	关键字	系统自提	
基金支持			

3.1.3 RDA 元素表的要素分析

RDA的重要特性之一是，继承了FRBR实体的聚集与组成。即它的模型结构能表现视为一个整体单元的实体，也能用同样的方法表现聚集与组成实体。从逻辑的观点来看，在会议资源中，它即可以代表由编者或者编纂者以文集/录形式汇集在一起的单篇会议文献的聚集，也可以表示一个会议录的一篇文章等。它们可以用同样的术语定义，具有同样的特征，而且以同样的方式互相联系。

在这种思路下，就可把原有的会议资源范围扩充到网络电子资源以外的形式，如图书、期刊、商业电子数据库等。基于这个的理解，为会议资源按照RDA的模型结构，进行第一组实体、资源关系的分析。

(1) 实体作品

“作品”是一个抽象的概念，即独有的知识或创作的艺术，在RDA中要求能够反应同一内容变化，能被认定。

重要会议资源中包括了三种资源类型，会议、会议录、会议文献。

(2) 内容表达

资源的“内容表达”可以分为文字表达、图像表达和声音表达。目前“重要开放会议资源采集与服务系统”和相关的馆藏会议资源，以文字表达方式为主，其中会议的介词的变化、界次的变化和缩略词的使用不能算作一个新的资源。

(3) 载体表现

原有的DC元素设计表主要表述的是网络资源的不同类型和格式，而重要会议服务与采集系统中的会议资源有部分也会参与到商业数据库、图书的出版。因此，在设计时要覆盖与网络会议资源相关联的其它载体。

载体形式上应该包括电子资源、图书、网络资源、期刊等。会议资源的“载体的表现”从中国科学院国家科学图书馆成都分馆现有的资源收藏情况来看，可以分为印本馆藏、期刊、电子期刊、CD-ROM、网络资源，集成网络资源（包括各类文摘库、全文库等）。

(4) 单件

在单件层面，即册、套、件被收藏的物件信息。目前以本馆馆藏为主，但在设计时要考虑未来与其他馆的关联。

如下是作品、内容表达、载体表现和单件的实例：

W1 会议录：Workshop on Agricultural Air Quality: State of the Science

E1 会议录的文字表达

M1 2006年以开放获取方式存放在网络上的文集（网络资源）

I1 中国科学院国家科学图书馆成都分馆馆藏

(5) 资源关系分析

从用户对资源使用需求的角度上来分析,对资源关系的需求可以分为三种类型^[25]:同一作品下的,不同内容表达层之间;不同载体表现层之间;不同单件之间的关系,另一种为不同作品之间的关系。

以会议录为例,第一类型中,同一会议录不同内容表达之间的关系。不同“内容表达”共同存在于同一个会议录下,通过“会议录”实现连接;不同载体表现之间通过相同的“内容表达”连接;不同单件之间通过相同“载体表现”形式建立关系。如:

W2 : 10th International Workshop of the Multimedia Metadata Community on Semantic Multimedia Database Technologies

E2 : 会议录的文字表达

M2 : 网络资源

I1=2 : 中国科学院国家科学图书馆成都分馆馆藏

M3 : 印本资源

I3 : 中国科学院国家科学图书馆成都分馆馆藏

I4 : 中国科学院计算机研究所馆藏

E3 : 会议文集图像表达

M4 : 网络资源

I5 : 中国科学院国家科学图书馆成都分馆馆藏

第二类型是不同作品之间的关系。这一部分主要通过不同作品之间的相同属性的关系建立联接。中国科学院国家科学图书馆的宋文老师对不同作品之间的常见关系²⁶,针对馆藏会议资源的特点,将资源之间的关系归纳为:

不同“会议文集”的相同责任者;

不同“会议文集”具有相同的主题,主题包括概念、地点、对象和事件;

不同“会议文集”之间的继承和被继承,合并与被合并等关系。

不同“会议文集”之间的“作者与主题”和“主题与作者”之间的关系。

第三种关系是部分和整体的关系。包括重要会议开放资源采集与服务平台中的一个页面、一篇论文的关系,馆藏丛书与分册的关系等。在重要会议开放资源采集与服务平台中,资源之间的联系可以通过会议资源的URL、会议题名、会议文献名、个人和团体、出版者、举办国、举办时间、会议主题建立联系。

按照以上分析,以会议录为列可以描述为:每个会议录通过视频、图片、声音、文字等形式(Expression)进行表达,每种形式又会通过如印本期刊、电子

²⁶宋文, FRBR 框架与信息资源的集成揭示[EB/OL]. [2012-03-15]
<http://ir.las.ac.cn/handle/12502/958?mode=full>

期刊、海报、录音、录像、CD-ROM等载体表达出来，这种物理体现构成了载体表现。一个载体的一种表现的样本为单件，如某会议的馆藏会议录。如图所示：

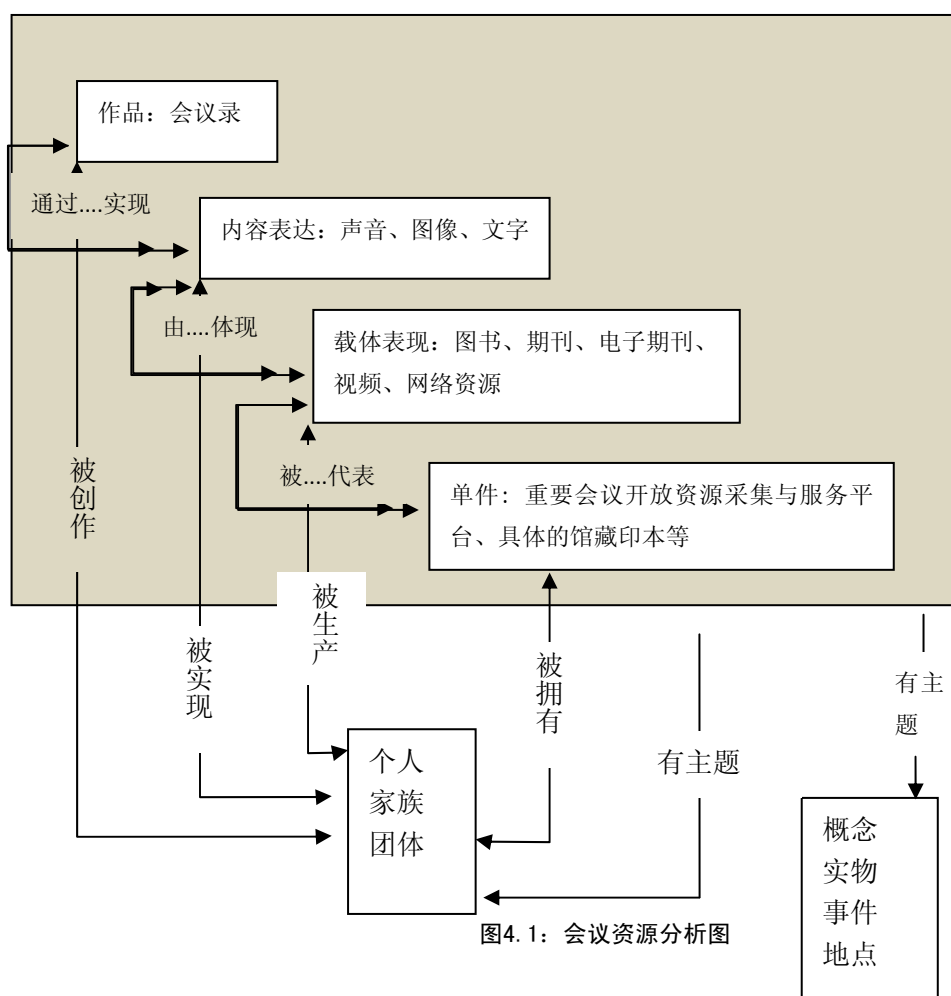


图4.1：会议资源分析图

3.1.4 RDA 元素表设计

首先，RDA元素设计时首先需要确定的是实体“作品”的选择。“作品”要能够揭示所描述会议资源的内容，能够与其它实体建立关联。

其次，对载体的覆盖。虽然RDA中载体不是设计的主要方面，但本系统中涉及的会议资源不仅是网络资源，还包括了印本、CD-ROM等载体，在原有的DC元素设计中没有考虑这类载体的内容反应，几种载体之间没有建立关联。

第三，关系的表达。本表设计的目的之一是为重要会议开放资源服务及采集系统能较DC元素更好的完成关系的关联，满足用户检索任务。即在描述时注意个人、家族和团体与作品的表达；作品与地点、概念、个人和团体的关系。

在设计时，应遵从以下原则：

应能对一个会议实体进行不同层次的独立描述；能较好的表达各会议实体资源之间的关系；能准确描述资源的属性；能考虑到与其它载体资源的联系；能被

加以修正以适用于不同的元数据环境；在概念术语和应用的范围等方面保持明确；与国际上通行的已经确立的原则模型标准保持兼容。

3.1.5 基于 RDA 的重要会议资源元素设计框架及元素集

3.1.5.1 基于RDA的重要会议资源元素设计框架

对于“重要会议资源采集与服务系统”来说，满足用户发现、识别、获取、选择是重点，RDA的元素在设计上侧重于会议资源的检索点、会议资源载体形式。按照作品的四个层级和三组实体将会议资源采取从上到下的设计，包括规范、书目、馆藏三个层次四个元素集。其逻辑关系如下：

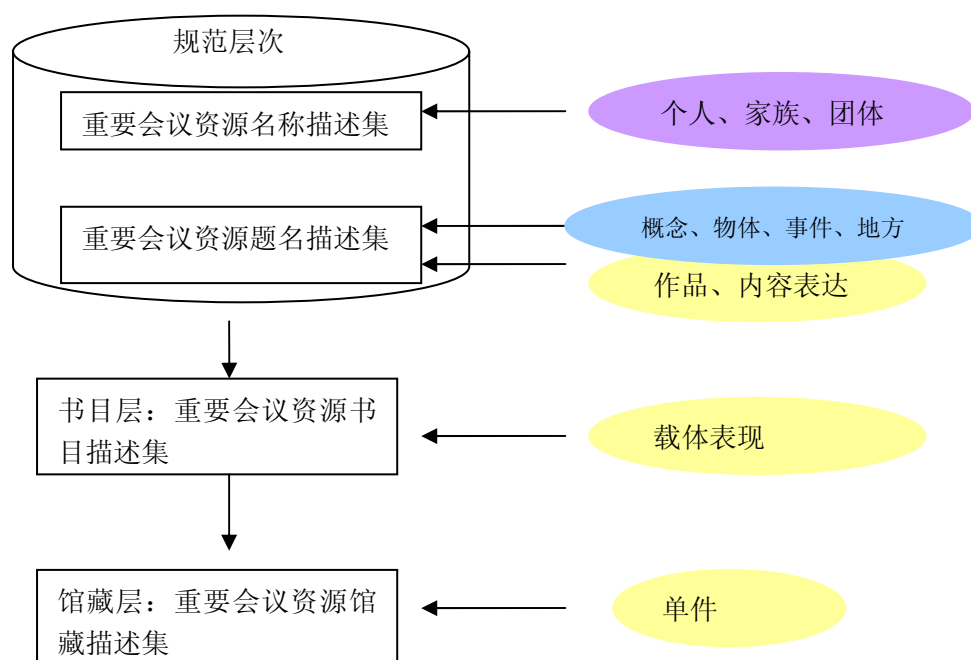


图3.2：会议资源元素集关系图

(1) 规范层次。针对作品和作品的表达方式，通过规范记录来描述作品的主题内容、与作品相关的学科、与作品相关的名称(个人、家族、团体)、作品的各种内容版本以及相关的作品等。在重要会议资源中，作品之间的关系和作品。作品之间的关系。包括部分和整体，包括章节、段落和分部，丛编中的单行出版物；“拥有”和“是...的部分”，在会议资源中体现出来就是会议录和单独的会议文献等。

(2) 书目层次,即不同载体。“重要开放会议资源采集与服务系统”的会议资源有基于载体的版本,如印本、数据库和其他光媒介质的版本;基于内容篇幅和语种的版本,如会议录、会议文献的不同语种版本在著录时需要做新增资源处理,会议录和会议文献摘要在处理时,应该形成一条新增资源。除此外,不同语种的会议文献也要新增一条记录。这些新增记录都对应应在相应的“内容表达”下。

(3) 馆藏层次的描述元素集。不作为描述的重点,但要考虑到未来与其他系统和图书馆的接口。

3.1.5.2 重要会议资源RDA元素设计表

(1) 会议资源名称规范

表3.4: 会议资源名称规范

RDA描述元素	域	会议资源相关要素	RDA元素等级
Preferred Name for the Corporate Body	corporate body	会议录团体作者	C-E
Preferred name for the person	person	会议文献作者	C-E
Address of the Person	person	会议文献作者联系地址	
		E-mail	
Preferred name for the corporate body	corporate body	会议主办者	C-E
Associated Institution	corporate body	会议相关机构	C-E
		会议作者相关机构	
Location of conference, etc.	corporate body	会议地址	C-E
Date of conference, etc.	corporate body	会议日期	C-E
Number of a conference, etc.	corporate body	会议号	C-E

(2) 会议资源题名规范

表3.5: 会议资源题名规范

描述元素	域	会议资源相关要素	元素等级
Creator	Work	会议举办者	C-E
		会议录作者	
		会议文献作者	

Other person, family, or corporate body associated with a work	Work	其他会议举办者	LCPS
		其他会议录作者	
		其他会议文献作者	
Summarization of the Content	expression	会议内容摘要	LCPS
Form of work	Work	相关会议	C-E
		卫星会议	
Related concept	-	会议相关主题	
		会议录相关主题	
		会议文献相关主题	
Date of work	Work	会议日期	C-E
Language of Expression	expression	会议录语种	C-E
		会议文献语种	
Content type	expression	会议文献类型	C-E
Numbering of Part	work	会议连续出版物 (IsPartOfSeries)	LCPS
Place of Origin of the Work	Work	会议地理等级范围	C-E

(3) 会议资源书目规范

表3.6: 会议资源书目规范

描述元素	域	会议资源相关要素	元素等级
Title proper	manifestation	会议题名	C-E
		会议录题名	
		会议文献题名	
Parallel title proper	manifestation	会议录其它语种题名	C-E
		会议文献其他语种题名	
Abbreviated title	manifestation	会议缩写题名	C-E
		会议录缩写题名	
		会议文献缩写题名	
Earlier title proper	manifestation	会议早期题名	LCPS

Later title proper	manifestation	会议最新题名	LCPS
Identifier for the Manifestation	manifestation	表现形式标识符（会议录号）	C-E
Uniform Resource Locator (URL)	-	会议 URL	LCPS
		会议录 URL	
		会议文献 URL	
Note on Publication Statement	manifestation	会议录摘要	LCPS
		会议录摘要	
		会议文献摘要	
		会议文献系统摘要	
Note on frequency	manifestation	会议频率	
		会议录发布频率	
Carrier Type	manifestation	会议录载体类型	C-E
		会议文献载体类型	
Media Type	manifestation	会议录格式	LCPS
		会议文献格式	
Date of production	manifestation	会议录发布日期	C-E
		会议文献发布日期	
Date of publication	manifestation	会议录发布日期	C-E
		会议文献发布日期	
Related Work	-	会议 (Is Part of)	C-E
		会议录名/会议文集名 (Is Part Of)	C-E
		会议 (Has Part Of)	C-E
		会议录名/会议文集名 (Has Part Of)	C-E
		连续出版物 (Numbering of Part)	LCPS
Place of publication	manifestation	会议录出版地	LCPS
		会议文献出版地	
Publisher's name	manifestation	会议录出版者	C-E
Date of publication	manifestation	会议录出版时间	
Media Type	manifestation	会议录格式	LCPS

		会议文献格式	
Carrier Type	manifestation	会议录载体	C-E
Note on Statement of Responsibility	manifestation	会议录权限申明	
Numbering of serials	manifestation	会议录卷期号	C-E
Terms of availability	manifestation	会议录获得方式	
		会议文献获得方式	
Related manifestation	manifestation	相关会议录	LCPS
		相关会议文献	
Related item	manifestation	相关会议录馆藏	LCPS
		相关会议文献馆藏	
Restrictions on access	-	会议录访问限制	
		会议文献访问限制	
Restrictions on use	-	会议录使用限制	LCPS
		会议文献使用限制	

(4) 会议资源馆藏规范

表3. 7: 会议资源馆藏规范

描述元素	域	会议资源相关要素	元素等级
Item-specific carrier characteristic	item	会议录载体特征	
		会议文献载体特征	
Item-specific carrier characteristic of early printed resource	item	早期印刷资源	
Identifier for the item	item	单件标识符	
Extent of Text	-	会议录页数	C-E/LCPS
		会议文献页数	

3.2 基于会议资源的 DC 描述

2009 年，作者在参与重要会议开放资源采集与服务系统研建过程中，按照会议资源按其资源特性，对该系统拟采集的资源进行了分析，确定了会议、会议录和会议文献三个资源描述对象，并依据 DC 的描述方式，建立了《重要开放会议元素名称、释义、描述规范》、《重要开放会议及会议录元素名称、释义、描述

规范》、《重要开放会议文献元素名称、释义、描述规范》三个元数据集。

在本次资源描述分析过程中，发现原有数据集中存在着一些不必要的元素，在会议和会议录资源的关联上也存在着问题。如，原来收集的资源主要是 PDF 格式的，在设计时没有考虑到其它格式，原来的会议录号也没有作为描述元素收入元素集；在关系的表述上，原来的会议资源和会议录作为两个分开的元素集进行描述，这样不利于资源关系的体现。故在原有研究基础上，对会议资源 DC 描述的内容进行了重新组织，将原有三表的内容划分为了两表，即“重要会议及会议录元素描述集”和“重要会议文献元素描述集”。

3.2.1 重要会议及会议录描述集

表 4.8: 重要会议及会议录元素描述集

核心集元素	与 DC 元素对映中文	与 DC 元素对映英文	元素修饰词
题名	名称	Title	会议正式名称
			会议其他名称
			会议缩写名称
			会议早期题名
			会议最新题名
			会议其他语种题名
			会议录题名
			会议录其他名称
主要责任者	创建者	Creator	会议主办者
			会议录团体作者
主题	主题	Subject and Keywords	表现形式标识符（会议录号）
			会议文献分类号
附注	描述	Description	会议机构描述
			会议摘要
			会议录摘要
出版者	出版者	Publisher	出版者
			会议录出版者

其他责任者	其他责任者	Contributor	相关机构
其他责任者	其他责任者	Contributor	相关机构
时间	日期	Data	首次会议举办时间
			会议日期
格式	格式	Format	会议录献格式/会议录载体特征
资源类型	类型	Type	会议分类
			会议录分类
资源标识	标识符	Identifier	会议 URL
			会议录 URL
来源	来源	Source	会议地址
语种	语种	Language	会议语种
			会议录语种
相关会议	关联	Relation	相关会议
			卫星会议 (Has Part Of)
			相关会议录/会议文集
			相关会议录
			会议举办频次
			会议已举办界次
相关会议录格式			
时空范围	时空范围	Coverage	会议地理等级范围
权限管理	权限	Rights management	会议录权限申明

3.2.2 重要会议文献描述集

表 3.9: 重要会议文献元素描述集

核心集元素	与 DC 元素对映中文	与 DC 元素对映中文	会议资源元素修饰词
题名	名称	Title	题名
			其它语种题名
			缩写提名

主要责任者	创建者	Creator	会议文献作者
			会议文献作者机构
			会议文献作者 Email
主题	主题	Subject Keywords and	会议文献关键词
			会议文献自由关键词
			会议文献分类号
附注	描述	Description	会议文献摘要
			会议文献系统摘要
			会议文献发布版本
出版者	出版者	Publisher	出版地
其他责任者	其他责任者	Contributor	会议文献其它责任者
时间	日期	Data	会议文献发布日期
			会议文献出版日期
资源类型	类型	Type	会议文献类型
格式	格式	Format	会议文献格式/会议文献载体特征
资源标识	标识符	Identifier	会议文献地址
来源	来源	Source	会议文献页数
语种	语种	Language	会议文献语种
相关	关联	Relation	会议 (Is Part Of)
			会议录名/会议文集名 (Is Part Of)
			会议连续出版物 (Is Part Of Series)
			会议 (Has Part Of)
			会议录名/会议文集名 (Has Part Of)
时空范围	时空范围	Coverage	内容覆盖范围
权限管理	权限	Rights	会议文献访问限制

		management	会议文献使用限制
--	--	------------	----------

3.3 会议资源 DC 到 RDA 的映射

3.3.1 映射原则

映射有多种级别的映射，本文中所述的映射主要是不同元数据之间元素间的映射，是语义级的映射。RDA和DC做为异构元数据规范，元素之间不可能做到精确对等的转换，但又需要对其映射做出质量上的规定。因此，在映射时，要遵从以下原则：

(1) 语义相似性原则

两种元数据间追求语义相似性，而不是术语的一致性。从元素的语义角度上的匹配，对其元素名称不要求完全对等。

(2) 最精确的匹配原则

即指在映射过程中，将两种元素尽可能的转换到最小粒度的术语单元。以RDA和DC的映射为例，应尽量将RDA的核心元素、元素下位类型和下位元素与DC的元素术语转换精确到DC元数据的修饰词层。

(3) 兼容性原则

当DC扩展元素中的一些内容在RDA中没有对应的元素进行映射时，可以将其转换到RDA的其它元素中。兼容性原则即将不同结构的元数据尽量的投射到相同语义内含的元素中去。

3.3.2 映射目标

在映射过程中，需要满足用户需求，可以揭示资源情况，实现以下目标：能够包含各种类型的资源和各种在目录或类似工具中表现的内容；不同的内容资源格式保持一致；能够容纳印刷格式或者超链视频的数字格式等；能实现对商业资源和开放资源的集成；能够集成物理馆藏和虚拟馆藏资源；在与RDA对应的元数据规范中找到相应的对应实体；能够实现对RDA和DC两种数据的互操作。

3.3.3 会议资源映射方法

对重要开放会议采集与服务系统的基于DC元数据的《重要开放会议元素名称、释义、描述规范》、《重要开放会议及会议录元素名称、释义、描述规范》、《重要开放会议文献元素名称、释义、描述规范》三个元素集进行分析和修改，将三个元素集合并为两个，描述对象为重要开放会议和会议录，重要开放会议文献。并参照《DC到RDA的映射表》的元素映射情况进行参照映射，最后形成《重要会议及会议录元素的DC到RDA映射表》、《重要会议文献元素的DC到RDA映射表》。

其中，RDA元素，按照RDA Toolkit正文里对元素的标注分为三类：核心元素（CORE ELEMENT，在下表中标识为C-E），国会图书馆推荐核心元素（CORE ELEMENT FOR Library Congress，下表中表述为LCPS）以及普通元素（下表中未做标示）。

3.3.4 会议资源 DC 到 RDA 的映射

映射情况如下：

（1）《重要会议及会议录元素的DC到RDA映射表》

本映射表在 DC 核心元素集与会议元素集、会议和会议录元素集的基础上，制定了相应的 RDA 的映射元素集，并根据 RDA 描述特征在 Title 部分增加了早期题名、会议最新题名、会议其他语种题名。同时，增加了“关系”的描述：相关作品、相关会议、下级系列会议。如下表。

表 3. 10: 重要会议及会议录元素的 DC 到 RDA 映射表

核心集元素	与 DC 元素对映中文	与 DC 元素对映英文	元素修饰词	RDA 元素	RDA 元素等级
题名	名称	title	会议正式名称	Title proper	C-E
			会议其他名称	Parallel title proper	
			会议缩写名称	Abbreviated title	
			会议早期题名	Earlier title proper	
			会议最新题名	Later title proper	
			会议缩写名称	Abbreviated title	
			会议录题名	Title proper	C-E
			会议录其他题名	Parallel title proper	
			会议录缩写题名	Abbreviated title	
主要责任者	创建者	creator	会议主办者	Creator	C-E
			会议录团体作者	Preferred Name for the Corporate Body	C-E
主题	主题	Subject and Keywords	表现形式标识符 (会议录号)	Identifier for the Manifestation	C-E
			会议分类	Identifier for the item	
			会议录分类	Identifier for the item	
			会议相关主题	Related concept	
			会议录相关主题	Related concept	
			会议关键词	N/A	
			会议录关键词	N/A	

			会议自由关键词	N/A	
			会议录自由关键词	N/A	
附注	描述	Description	会议机构描述	Associated Institution	C-E
			会议摘要	Summarization of the Content	LCPS
			会议录摘要	Note on Publication Statement	LCPS
出版者	出版者	Publisher	出版者	Publisher's name	C-E
			会议录出版者	Publisher's name	
其他责任者	其他责任者	Contributor	相关机构	Associated Institution	C-E
时间	日期	Data	首次会议举办时间	Date of Work	C-E
			会议日期	Date of Conference	C-E
			会议录发布日期	Date of production	C-E
			会议录出版日期	Date of publication	C-E
资源类型	类型	Type	会议格式	Content Type	C-E
			会议录格式	Content Type	C-E
格式	格式	Format	会议媒体类型	Media Type	LCPS
			会议录美的类型		
			会议载体版本	Carrier Type	C-E
			会议录载体版本		
资源标识	标识符	identifier	会议 URL	Uniform Resource Locator (URL)	LCPS
			会议录 URL	Uniform Resource Locator (URL)	
来源	来源	source	会议地址	Location of Conference, etc.	C-E
语种	语种	language	会议录语种	Language of Expression	C-E
相关会议	关联	relation	相关会议	Form of work	C-E
			卫星会议 (has part of)	Form of work	C-E
			相关会议录	Related Work	C-E
			会议举办频次	Note on frequency	
			会议号	Number of a Conference, etc.	C-E

			相关会议录格式	Related Manifestation	C-E
时空范围	时空范围	coverage	会议地理等级范围	Place of Origin of the Work	C-E
权限管理	权限	Rights management	会议录访问限制	Restrictions on access	
			会议录使用限制	Restrictions on use	LCPS

(2) 《重要会议文献元素的 DC 到 RDA 映射表》

参考 DC 核心元素，制定了对应的 RDA 会议资源元素，如下表。

表 3.11: 重要会议文献元素的 DC 到 RDA 映射表

核心集元素	与 DC 元素对映中文	与 DC 元素对映中文	会议资源元素修饰词	RDA 元素	元素等级
题名	名称	Title	会议文献题名	Title proper	C-E
			会议文献其它题名	Parallel title proper	C-E
			缩写题名	Abbreviated title	C-E
主要责任者	创建者	Creator	会议文献作者	Preferred name for the person	C-E
			会议文献作者机构	Associated Institution	C-E
			会议文献作者 email	Address of the Person	
主题	主题	Subject and Keywords	会议文献分类号	Identifier for the item	
			会议文献相关主题	Related concept	
			会议文献关键词	N/A	
			会议文献自由关键词	N/A	
附注	描述	Description	会议文献摘要	Note on Publication Statement	LCPS
			会议文献系统摘要	Note on Publication Statement	
出版者	出版者	Publisher	会议文献出版地	Place of Publication	LCPS
其他责任者	其他责任者	Contributor	会议文献其它责任者	Other person, family, or	LCPS

				corporate body associated with a work	
时间	日期	Data	会议文献发布日期	Date of production	C-E
			会议文献出版日期	Date of Publication	C-E
资源类型	类型	Type	会议文献内容类型	Content Type	C-E
格式	格式	Format	会议文献格式	Media Type	
				Carrier Type	C-E
资源标识	标识符	Identifier	会议文献 URL 地址	Uniform Resource Locator	LCPS
来源	来源	Source	会议文献页数	Extent of Text	C-E/LCPS
语种	语种	Language	会议文献语种	Language of Expression	C-E
相关	关联	Relation	会议 (is part of)	Related Work	C-E
			会议录名/会议文集名 (is part of)	Related Work	C-E
			会议录卷期号 (IsPartOfSeries)	Numbering of Part	LCPS
			会议 (has part of)	Related Work	C-E
			会议录名/会议文集名 (has part of)	Related Work	C-E
权限管理	权限	Rights management	会议文献访问限制	Restrictions on Access	
			会议文献使用限制	Restrictions on use	LCPS

3.4 RDA 与 DC 在描述资源中的效用比较

(1) 语义粒度差异

DC 与 RDA 的语义粒度相比,不够细。在映射过程中,我们发现:DC 的 Identifier 主要包括网络资源的标识词和书目资料的标识词,分类较粗。而 RDA 将标识词分别在两组实体中得到体现,标识词的分类更详尽。

RDA 中的 Identifier 在第一组实体中分别有体现:

RDA :2.15 Identifier for the Manifestation

RDA: 6.13 Identifier for the Expression

RDA: 6.8 Identifier for the Work

RDA: 2.19 Identifier for the Item

在第二组实体中:

RDA: 29.4.1 Identifier for the Related Person, Family, or Corporate Body

RDA: 9.18 Identifier for the Person

RDA: 16.3 Identifier for the Place(尚无内容)

RDA: 11.12 Identifier for the Corporate Body

RDA: 10.9 Identifier for the Family

除此外，RDA 在 TITLE 和 DATA 中还涉及到官员的称呼、宗教的题名、法律文书方面的题名和签验时间等更为详细的内容。因此，DC 和 RDA 的出发点不同，决定了他们的资源描述粒度和内容不同。

(2) 不对应现象

DC 中的一些内容在 RDA 中暂时没有反应。如在主题分类上，RDA 没有对 Dewey Decimal call number, Local call numbers, 等做定义。只将 Classification number 对应到 Identifier for the Manifestation 上。

(3) 语义描述效果

DC 和 RDA 在映射的关系上存在的四种映射情况

一是一对一的映射关系。DC 元数据中的每一个元素在 RDA 中都存在着多元素的描述，一对一的映射关系从研究情况来看，目前还没有发现。

二是一对多的关系。DC 题名项中“[Alternative title](#)”对应到 RDA 中对不同题名的描述，如 Parallel Title Proper（并列题名）、Other Title Information（其它题名）、Parallel Other Title Information（并列其它题名）、Variant Title（变种题名）、Earlier Title Proper（早期题名）、Later Title Proper（最新题名）、Key Title（关键题名）和 Abbreviated Title(缩略题名)。除此外，还有元素“Format”可以对应到 RDA 中的“Carrier type”和“Media type”。

三是多对一的关系。如 DC “Publisher”和“Creator”元素对应到 RDA 元素中个人、团体和家族实体关系部分。

四是无对应关系，如 DC 中的“Subject”主题部分在 RDA 中没有对应的元素关系，RDA 对主题的描述主要通过概念、地点、对象和事件四个共同表达，

没有单独就主题列出相应元素。

(4) 兼容性和互操作

从 DC 元数据与 RDA 元数据映射分析、《重要会议及会议录元素的 DC 到 RDA 映射表》、《重要会议文献元素的 DC 到 RDA 映射表》三种映射情况分析来看，RDA 能够较准确的描述出 DC 元数据中的元素表达内容。DC 为自己的互操作水平确定了四个级别²⁷，本文从这四个互操作水平来分析二者情况，如下：

第一级，共享术语定义。

在这个水平层次上，互操作元数据使用基于自然语言的定义。在应用环境中，如互联网，图书馆系统或者联合数据库，参与同意在他们的元数据中使用和定义这些术语。从这个角度来看，目前大多数现有的元数据已经实现在这一级别的操作性。

第二级，语义互操作

DC 与语义互操作关联最大的即其推荐的编码方案之一就是 RDF。RDF 是用来支持与数据共享的正式模型，是元数据的应用程序的互操作性之间的基础。在维基百科的定义中，“关联数据”在语义网中描述的“推荐的最佳做法是使用 URI(Web addresses)和 RDF 公开，共享和连接件的数据，信息以及知识语义”²⁸。

近年来，DCMI 元数据术语的属性和类已经被定义为与关联数据原则的兼容性。这促进了 DC 在与非图书馆社区的应用如，商业和公共部门的数据已添加了越来越多的链接的数据云。RDFa(Resource Description Framework – in – attributes)的实施上，支持在网页中嵌入的链接数据的方法搜索引擎，如对雅虎和如 Drupal 内容管理平台的支持。事实上，都柏林核心成立之初的理念 —— “为实现资源发现的简单元数据”正在“结构化数据优化搜索引擎”的旗帜下改变。从这个角度来讲，DC 在这一个水平层上会出现较快的变化。

第三级，描述语法互操作性集。

应用程序与链接数据模型兼容，共享可验证元数据记录的抽象语法，即“描述集”。RDA 使用的模型来源于“实体-关系”模型，是 FR 家族的核心思想，从目前 FRBR 在各社区的应用情况来看，与其它应用程序和链接数据模型兼容度较好。在这一点上可与 DC 一较高下。

第四级，描述集资料的互通。

记录之间交换元数据的应用程序遵循，此外，有一套共同的约束，使用相同的词汇，反映了世界共享模型。3 级和 4 以上级别 1 和 2 因为他们都没有得到很好的支持软件工具的试验，但在这项工作中解决的问题预计将增长的重要

²⁷ DCMI, Levels of interoperability [EB/OL]. [2012-02-29] <http://dublincore.org/metadata-basics/>

²⁸ Linked Data. [EB/OL]. [2012-04-29] <http://linkeddata.org/>

性，他们的信息元数据记录的生产者移动到一个链接的数据环境。RDA 目前通过 XC 的支持，在元数据的互操作上，有较好的可操作性。

从这四个级别来说，RDA 和 DC 在“共享术语定义”级别上表现相当，“语义互操作”上，DC 较 RDA 更专一，但内容不如 RDA 丰富，这也意味着，RDA 操作起来更复杂。在“描述语法互操作”上各有特长，在“描述集资料的互通”方面，两者都有较好的可操作性。

（5）描述过程

RDA 元数据最大的特色之一即其描述“关系”的功能。在基于重要会议开放资源采集与服务系统的会议资源描述中，通过 RDA 实现了在资源的综合集成上的情况，即传统馆藏和数字资源的载体上的集成，弥补了 DC 元数据以描述网络资源为特色的描述方式。

其次，在反应用户需求上，DC 的每一个元素都可以作为检索点。RDA 则通过强调几组实体之间的关系来实现用户检索需求，元素与元素间的关联度较 DC 元数据强，但在描述过程中需要对用户的检索需求进行分析，操作起来较 DC 元数据复杂。

第三，从 RDA 的语义描述情况来看，RDA 因为继承了 AACR2 的描述方式，故在资源的细节描述上占有优势，但描述元素较多，要作到精确选择需要一定的培训。

（6）使用状况（编目人员）

DC 元数据是从简到繁，RDA 是想要面面俱到。从编目人员的角度来看，RDA 更能描述资源的各个特征，著录的内容覆盖面较广；相应的 DC 格式简单，使用方便，可以扩充内容，但对资源的描述内容有所限制，传统载体的资源不在其研究范围内，对图书馆编目人员来说，同时使用多种描述方工描述资源不方便。从其它社区应用人员的角度来说，RDA 内容广博，但需要熟练的掌握需要培训和时间；而 DC 结构简单，使用人员无需培训，目前在其它社区的应用上已经有较好的效果。

从应用的便捷性上来说，DC 的元素表设计是根据需求来进行的设计，RDA 的设计是一个大而全的标准，对各项可能存在的描述都进行了规范和说明，在实际应用中，需要使用人员去掉不必要的元素的干扰。